

Homework 1

Mark Schulist

2024-08-30

1.2.4

- a. The population is air in airplanes on domestic flights.
- b. The sample is the 175 air samples from 175 flights.
- c. The characteristic of interest is the “degree of staleness” of the air from the air quality tests.

1.3.2

No, we have stratified sampling. Simple random sampling would mean that we are randomly drawing from the ENTIRE population, not randomly drawing (proportionally) from each strata.

1.4.4

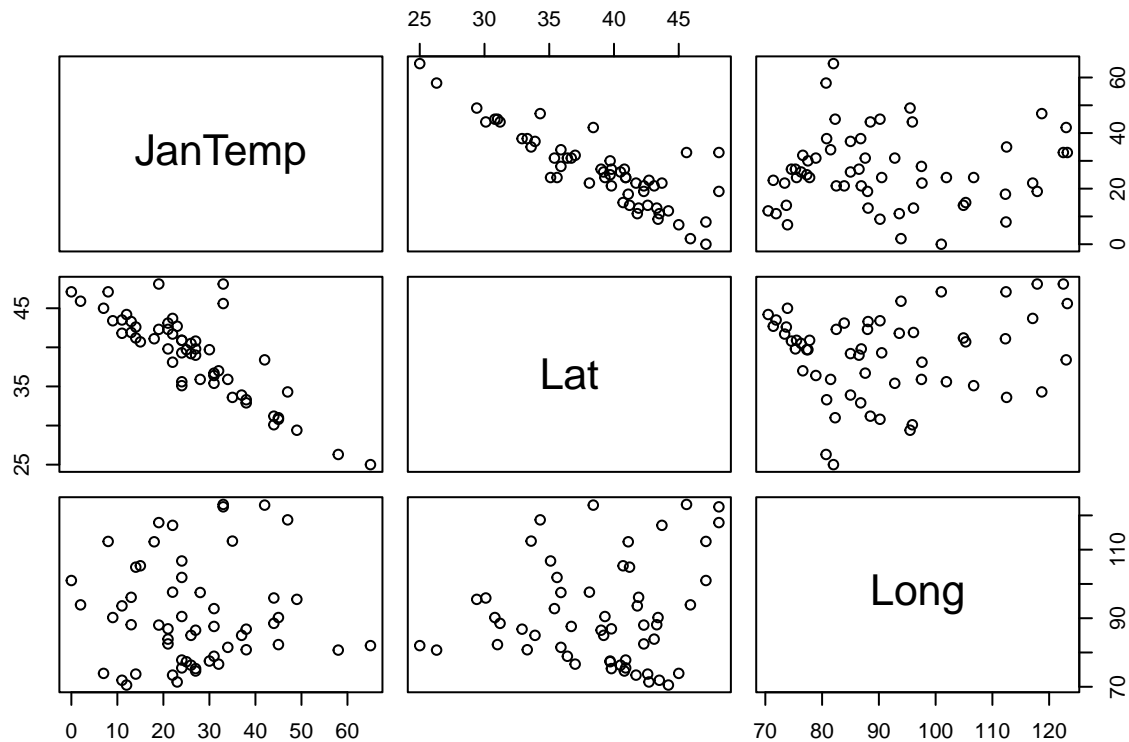
- a. The variable of interest is the “degree of staleness.” The statistical population is the staleness on every domestic flight.
- b. Quantitative, the “degree” is measured with a device numerically.
- c. Univariate, we only have one variable (staleness).

1.5.4a

a.

```
url <- "https://media.pearsoncmg.com/cmg/pmmg_mml_shared/mathstatsresources/Akritis/Temp.Long.Lat.txt"
temp <- read.table(url, header = T)

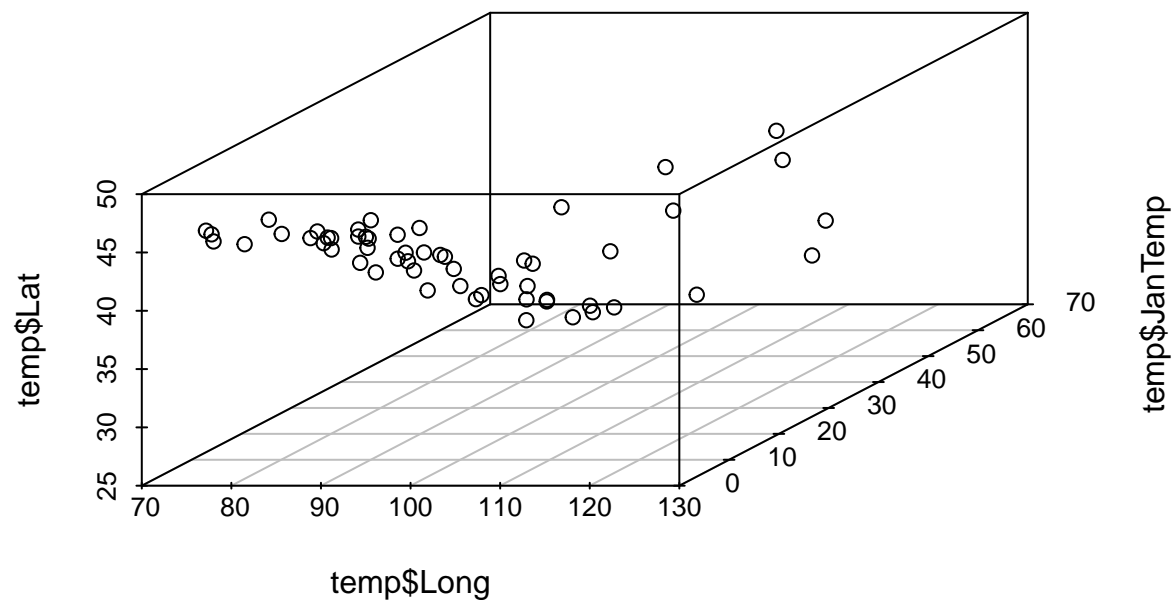
pairs(temp[,2:4])
```



Latitude seems to be a better predictor as there is a clear (negative) relationship between JanTemp and Lat. There does not appear to be any relationship between JanTemp and Long.

b.

```
library(scatterplot3d)
scatterplot3d(temp$Long, temp$JanTemp, temp$Lat)
```



It is challenging to see in the plot because 3 dimensions is difficult to comprehend. That being said, we can see the same trends as above, where Lat and Temp are very strongly correlated and Long and Temp are not correlated very much. We can see that Lat forms a nice line, whereas Long is much more scattered.

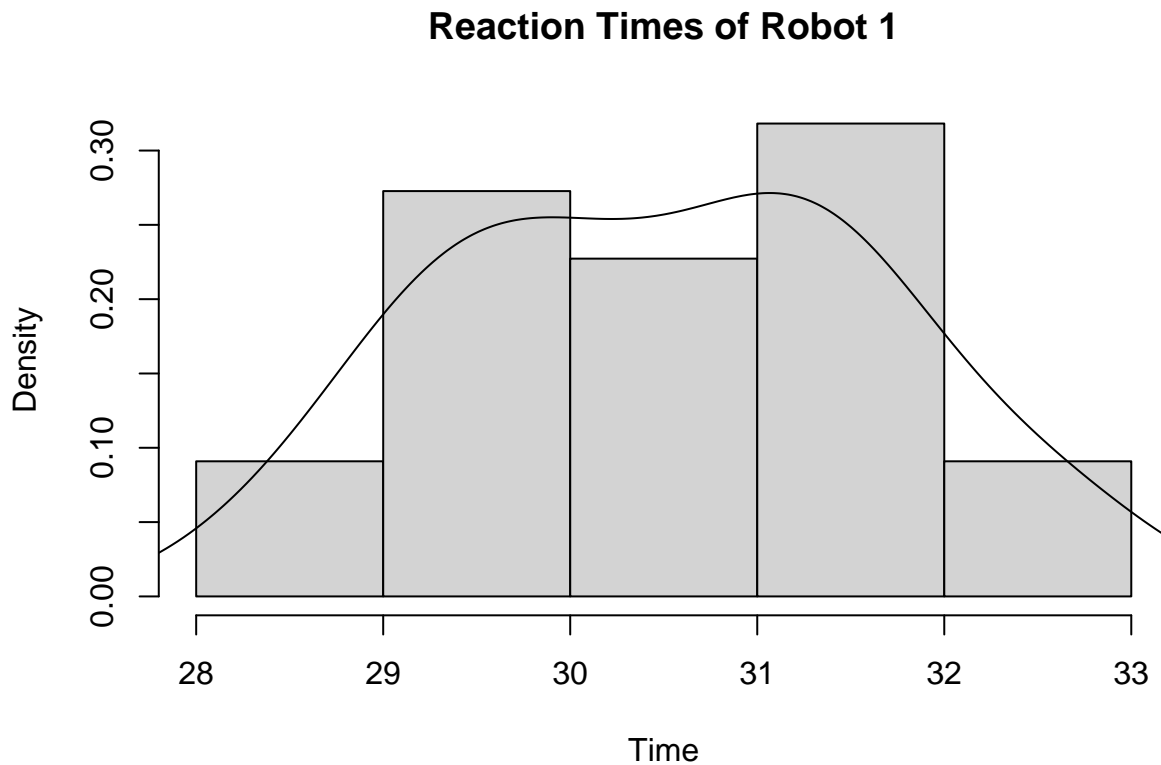
1.5.9

```
url <- "https://media.pearsoncmg.com/cmg/pmmg_mml_shared/mathstatsresources/Akritis/RobotReactTime.txt"
t <- read.table(url, header = T)

attach(t)
t1 <- Time[Robot==1]
```

a.

```
hist(t1, freq = F, main = "Reaction Times of Robot 1", xlab = "Time")
lines(density(t1))
```



b.

```
stem(t1)
```

```
##
## The decimal point is at the |
##
## 28 | 4
## 29 | 0133688
## 30 | 03388
## 31 | 0234669
## 32 | 47
```